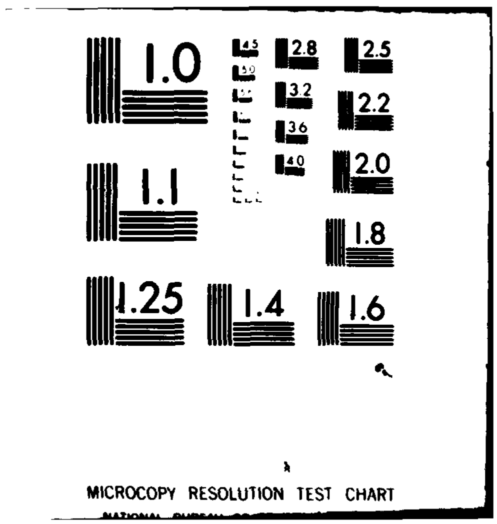
AD-A097 460 PRINCETON UNIV NJ DEPT OF STATISTICS F/6 12/1
ANTIPODALLY SYMMETRIC DISTRIBUTIONS FOR ORIENTATION STATISTICS.(U)
MAR 81 M J PRENTICE N00014-79-C-0322
UNCLASSIFIED TR-192-SER-2 NL

1 of 1

END
DATE
FILMED
5-81
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
MICROCOPY RESOLUTION TEST CHART

LEVEL II

(12)

ADA097460

(6)

ANTIPODALLY SYMMETRIC DISTRIBUTIONS FOR ORIENTATION STATISTICS.

by

(10) Michael J. Prentice
Department of Statistics
Edinburgh University
Scotland, U.K.

(14) TR-192-SER-2

(9) Technical Report No. 192, Series 2
Department of Statistics
Princeton University

(11) Mar[redacted]81

(12) 17

(15) Research supported in part by a contract with the Office of Naval Research, No. N00014-79-C-0322 awarded to the Department of Statistics, Princeton University, Princeton, New Jersey.

406 873

81 4 7 04

Antipodally symmetric distributions for orientation statistics

by

Michael J. Prentice
Department of Statistics
Edinburgh University
Scotland, U.K.

S U M M A R Y

The conventional antipodally symmetric Bingham matrix distribution on the Stiefel manifold is generalised. Large sample maximum likelihood estimation and uniformity tests are discussed, and a parametric model for axial orientations (X-shapes) is suggested. A generalisation of the Khatri-Mardia matrix distribution is developed to provide a model suitable for hybrids (T-shapes). Beran's results on exponential models for directional data are extended to orientation statistics to provide regression estimators and goodness-of-fit tests as alternatives to maximum likelihood estimation and likelihood ratio tests.

Keywords: Orientation statistics, Bingham matrix distribution, antipodal symmetry, exponential family.

AMS 1970 Subject classifications. Primary 62F10, secondary 62F05.

1. Introduction

Following Downs (1972), we define an orientation statistic as a rigid m-frame in R^p $(m \leq p)$ i.e. an $m \times p$ matrix X s.t. $XX' = C$, where C is an $m \times m$ symmetric positive definite matrix specifying the angles between the rows of X. Without loss of generality we suppose $C = I_m$ since all methodology for $C = I_m$ can be extended trivially for general C. We define an L-shape to be a rigid m-frame of signed directions (a conventional orientation statistic), an X-shape to be a rigid m-frame of axes, and a T-shape to be a hybrid, an m-frame of m_1 axes and $m_2 = m - m_1$ signed directions in R^p. All methodology so far published is suitable for L-shapes. We assume that $p \geq 3$ since when $p = 2$ every type of orientation statistic reduces to either a direction or an axis on the circle.

The von Mises-Fisher matrix distribution provides a suitable unimodal pdf. for orientation statistics. Maximum likelihood estimates and likelihood ratio tests have been developed by Downs (1972), Khatri and Mardia (1977) and Jupp and Mardia (1979). The conventional fully parameterised Bingham matrix distribution (Khatri and Mardia, 1977, (7.2) with $\underline{\delta} = \underline{0}$) is the obvious analogue on the Stiefel manifold of Bingham's antipodally symmetric distribution on the sphere (Bingham, 1974). Maximum likelihood estimation and tests of randomness against a special case of this distribution have been treated by Jupp and Mardia (1979) and Mardia and Khatri (1977). In Section 2 we state the corresponding MLE results for the fully parameterised $2^{m'}$-modal Bingham matrix distribution. We extend these results to a generalisation of the Bingham matrix distribution in Section 3, and obtain a new 'Bingham' statistic as a large sample test of uniformity against general antipodally symmetric alternatives, a large sample approximation to the likelihood ratio statistic. In Section 4 a special case of this distribution is suggested as a suitable model for X-shapes, and an axial Bingham statistic is obtained for a test of uniformity. The generalised Khatri-Mardia matrix distribution is used in Section 5 to provide a suitable parametric model for T-shapes, and a hybrid Rayleigh-Bingham statistic is obtained as a large sample test of uniformity. Beran's results on rotationally invariant exponential models for directional data are extended to the Stiefel manifold in Section 6, to provide regression estimators and goodness of fit tests within the generalised Khatri-Mardia family of distributions, as alternatives to the computationally inconvenient maximum likelihood estimators and likelihood ratio tests.

2. The Bingham matrix distribution

If X is an $m \times p$ matrix random variable ($m \leq p$) with pdf

$$(2\pi)^{-\frac{1}{2}mp} \, |K|^{\frac{1}{2}p} \, |V|^{\frac{1}{2}m} \, \mathrm{etr}(-\tfrac{1}{2}K(X-\mu)V(X-\mu)')dX \qquad (2.1)$$

where K is $m \times m$ symmetric positive definite, and V is $p \times p$ symmetric positive definite, then if $\mu = 0_{mp}$, an $m \times p$ matrix of zeros, the pdf of X, conditional on $XX' = I_m$ is (Khatri and Mardia, 1977, (7.2))

$$\mathrm{etr}(-\gamma m^{-1} I_m - \tfrac{1}{2}KXVX') \, [dX] \qquad (2.2)$$

where $[dX]$ denotes the uniform distribution on the Stiefel manifold $O(m,p)$, and $\gamma = \gamma(K \otimes V)$ is a normalising constant which depends only upon the diagonal matrices D_K, D_V of eigenvalues of K and V. Series expansions for $e^{+\gamma}$, a hypergeometric function of two matrix arguments, have been given by Srivastava and Carter (1980). For reasons which will become apparent, we denote the distribution (2.2) $B(X,m,p,K \otimes V)$. Its parameter space has dimension $\binom{m+1}{2} + \binom{p+1}{2} - 2$ when $1 \leq m \leq p-1$, since $B(X,m,p,(\alpha K) \otimes (\alpha^{-1} V + \beta I_p)) = B(X,m,p,K \otimes V)$ for all real scalars $\alpha \neq 0, \beta$ (cf. Bingham, 1974, Lemma 2.1). Similarly, when $m = p$ $B(X,p,p,(\alpha K + \beta_1 I_p) \otimes (\alpha^{-1} V + \beta_2 I_p)) = B(X,p,p,K \otimes V)$, where without loss of generality we may assume trace(K-V) = 0 , so that the parameter space has dimension $2\binom{p+1}{2} - 3$. Where convenient to identify parameters uniquely we shall assume that $v_{pp} = 0$ (and $k_{pp} = 0$ if $m = p$), or, if dealing in spectral decompositions, that the entries in $-D_K$ and D_V are in decreasing order with $(D_V)_{pp} = 0$ (and $(D_K)_{11} = 0$ if $m = p$). We assume that the elements of D_K and D_V are distinct so that K has a unique matrix $Q \in O(m', m)$ of eigenvectors and V has a corresponding matrix $M \in O(p-1, p)$. From Theobald (1975, Theorem 1), the distribution (2.2) has $2^{m'}$ modes at the points $X = Q'M_1$, where M_1 is the $m' \times p$ matrix of the first $m' = \min(m, p-1)$ columns of M. The multiplicity $2^{m'}$ arises from the possible sign changes of the columns of Q and M_1 . I am grateful to Dr Theobald for first drawing my attention to this.

A random sample $X_1, X_2, \ldots, X_n$ on $O(m,p)$ from the distribution (2.2) has log likelihood

$$-n\gamma - \tfrac{1}{2}n \operatorname{trace}(K \otimes V)Y \tag{2.3}$$

where $Y = n^{-1} \sum_{i=1}^{n} X_i \otimes X_i' = (y_{jk,\ell q})$, $1 \le j,q \le m$, $1 \le k,\ell \le p$. The distinct elements of Y , excluding those for which $(k,\ell) = (p,p)$ (and also those for which $(j,q) = (p,p)$ when $m = p$) are sufficient but not minimally sufficient unless $m = 1$ (the case of directional data), because $v(m,p) = \binom{mp+1}{2} - \binom{m+1}{2} \ge \binom{m+1}{2} + \binom{p+1}{2} - 2$, with equality only when $m = 1$, and $v(p,p) = \binom{p^2+1}{2} - 2\binom{p+1}{2} + 1 > 2\binom{p+1}{2} - 3$. The following is a consequence of Berk (1972), and elementary calculus.

<u>Theorem 2.1</u>

(a) For sufficiently large n , there exist unique MLEs $\hat{K}, \hat{V}$ subject to the condition $\hat{v}_{pp} = 0$ (and $\hat{k}_{pp} = 0$ and $\operatorname{trace}(\hat{K}-\hat{V}) = 0$ when $m = p$) of the parameters in $B(X, m, p, K \otimes V)$. The MLEs are the solutions of

$$\left(\frac{\partial\gamma}{\partial K}\right)_{(K,V)=(\hat{K},\hat{V})} = -\tfrac{1}{2} W_1(\hat{V})$$

and

$$\left(\frac{\partial\gamma}{\partial V}\right)_{(K,V)=(\hat{K},\hat{V})} = -\tfrac{1}{2} W_2(\hat{K})$$

where $W_1(V) = n^{-1} \sum_{i=1}^{n} X_i V X_i'$ and $W_2(K) = n^{-1} \sum_{i=1}^{n} X_i' K X_i$.

(b) (Spectral version) If $K = Q'D_K Q$ and $V = M'D_V M$ are unique spectral decompositions of K and V, where $Q \in O(m',m)$, $M \in O(p-1,p)$ and D_K, D_V are respectively $m' \times m'$ and $(p-1) \times (p-1)$ diagonal matrices (and trace $(D_K - D_V) = 0$ if $m = p$), then for sufficiently large n with probability 1 there exist unique MLEs $\hat{Q}$, $\hat{M}$, $\hat{D}_K$, $\hat{D}_V$ of Q, M, D_K, D_V given by the unique spectral decompositions

$W_1(\hat{V}) = \hat{Q}'\hat{D}_{W_1}\hat{Q}$, $W_2(\hat{K}) = \hat{M}'\hat{D}_{W_2}\hat{M}$ and the equations

$$\left(\frac{\partial\gamma}{\partial D_K}\right)_{(D_K,D_V)=(\hat{D}_K,\hat{D}_V)} = -\tfrac{1}{2} \hat{D}_{W_1}$$

$$\left(\frac{\partial\gamma}{\partial D_V}\right)_{(D_K,D_V)=(\hat{D}_K,\hat{D}_V)} = -\tfrac{1}{2} \hat{D}_{W_2} \quad .$$

We offer no algorithm for the evaluation of $\hat{K}$ and $\hat{V}$. Given suitable initial approximations $\hat{K}_0$, $\hat{V}_0$ it should be possible to construct an iterative procedure, given tractable series expansions for γ and its first derivatives (see Bingham, 1976), but considerable computational effort is necessary.

Jupp and Mardia (1979, Theorems 1(b) and 4) have obtained the analogue of Theorem 2.1 for the special case $B(X,m,p,I_m \otimes V)$ which has parameter space of dimension $\tfrac{1}{2}(p-1)(p+2)$ provided $m < p$. We note that $B(X,p,p,I_p \otimes V)$ is the uniform distribution so their results are invalid for square orientations. The log likelihood of a sample from

$B(X,m,p,I_m \otimes V)$, $m<p$,

is $-n\,\gamma(I_m \otimes V) - \tfrac{1}{2}n$ trace VY^* (2.4)

where $Y^* = n^{-1} \sum_{i=1}^{n} X_i' X_i$ is $p \times p$ symmetric, positive definite with probability 1. The distinct elements of Y^*, excluding y^*_{pp} say, are minimally sufficient for V, subject to $v_{pp}=0$ say. A simple large sample test of uniformity against alternatives (2.4), asymptotically equivalent to the likelihood ratio statistic, may be obtained by generalising Bingham's Theorem 5.2 (1974, p.1208). We obtain

<u>Theorem 2.2</u>

On the null hypothesis of uniformity the statistic $S_u = (\text{trace}(Y^{*2}) - m^2/p)\, np(p-1)(p+2)/2m(p-m)$ is asymptotically distributed as χ^2 on $\nu(1,p) = \tfrac{1}{2}(p-1)(p+2)$ degrees of freedom.

<u>Proof</u> (A simpler version of Theorem 3.2 below).

If $\underline{y}^*$ is the $\binom{p+1}{2}$-vector of the distinct elements of Y^*, then Σ^*, the asymptotic variance matrix (Mardia and Khatri, 1977 p.469) of $n^{\frac{1}{2}}\underline{y}^*$ has rank $\nu(1,p)$ and a generalised inverse $[p(p-1)(p+2)/2m(p-m)]$ blockdiag $(I_p, 2I_{\binom{p}{2}})$. Since Y^* has null expectation $mp^{-1} I_p$, the result follows immediately (see also Khatri and Mardia, 1977, p.471, for an alternative derivation).

<u>Remarks</u> 1. S_u is undefined when $m=p$

2. $B(X,m,p,K \otimes I_p)$, the other obvious special case of (2.2), is the uniform distribution for all m, $1 \le m \le p$.

3. <u>A generalisation</u>

Consider the fully parameterised multivariate normal model

$$(2\pi)^{-\frac{1}{2}mp}\, |A|^{\frac{1}{2}mp}\, \text{etr}(-\tfrac{1}{2}\, A(X-\mu) \otimes (X-\mu)')dX \qquad (3.1)$$

and the conditional distribution

$$\mathrm{etr}(-\gamma(mp)^{-1} I_{mp} - \tfrac{1}{2} AX \otimes X')[dX] \qquad (3.2)$$

obtained when $XX' = I_m$ and $\underline{\mu} = \underline{0}_{mp}$. Here A is $mp \times mp$ symmetric positive definite with distinct eigenvalues $a_1 > \ldots > a_{mp}$, and $\gamma = \gamma(A)$ is a normalising constant. The distribution (3.2), denoted $B(X,m,p,A)$ is antipodally symmetric and has parameter space of dimension $\nu(m,p)$, since $B(X,m,p,A) = B(X,m,p,A + A_1 \otimes I_p)$ for all real $m\times m$ symmetric matrices A_1, and $B(X,p,p,A) = B(X,p,p,A + A_1 \otimes I_p + I_p \otimes A_2)$ for all real $p\times p$ symmetric matrices A_1, A_2, where without loss of generality we may assume $\mathrm{trace}(A_1 - A_2) = 0$. Where convenient to identify parameters uniquely we shall assume that $A = (a_{jk,\ell q})$ $1 \le j, q \le m$, $1 \le k, \ell \le p$, satisfies the conditions

$$a_{jp,pq} = 0 \text{ for all } j,q, \text{ and, if } m = p, \; a_{pk,\ell p} = 0 \text{ for all } k,\ell . \qquad (3.3)$$

A random sample from the distribution (3.2) has log likelihood

$$-n\gamma(A) - \tfrac{1}{2}n \,\mathrm{trace}\, AY \qquad (3.4)$$

with $Y = (y_{jk,\ell q})$ as in (2.3). The distinct elements of Y, excluding those corresponding to (3.3), are minimally sufficient for A. From Berk (1972), and differentiation we obtain.

<u>Theorem 3.1</u>

For sufficiently large n, with probability 1 there exists a unique MLE $\hat{A}$ of A subject to (3.3), which is the solution of

$$\left(\frac{\partial\gamma}{\partial A}\right)_{(A=\hat{A})} = -\tfrac{1}{2}Y .$$

As with theorem 2.1, there are considerable computational difficulties associated with the search for $\hat{A}$. In particular, no explicit form of $\gamma(A)$ is currently available.

A Bingham statistic for testing uniformity against all antipodally symmetric alternatives of the form $B(X,m,p,A)$ may be obtained as follows.

Theorem 3.2

Let $Z_1 = Y - p^{-1} I_{mp}$ and $Z_2 = (Z^{(2)}_{jk,\ell q}) = \frac{1}{2}(Y_{jk,\ell q} - Y_{qk,\ell j})$. On the null hypothesis of uniformity the statistic

$B_{mp} = n(p-1)[\frac{1}{2}(p+2) \text{ trace } (Z_1'Z_1) - \text{trace } (Z_2'Z_2)]$ is asymptotically distributed as χ^2 on $\nu(m,p)$ degrees of freedom.

Proof

On the null hypothesis Y has expectation $p^{-1} I_{mp}$. Consider first the case $m<p$. The $\binom{mp+1}{2}$ distinct elements of Y may be written as a vector

$$\underline{y} = (<\underline{y}_{11}>, \ldots, <\underline{y}_{mm}>, <\underline{y}_{12}>, \ldots, <\underline{y}_{m-1,m}>)'$$

where each $<\underline{y}_{jj}>$, $1 \le j \le m$, is a $\binom{p+1}{2}$-vector

$$(y_{j1,1j}, y_{j2,2j}, \ldots, y_{jj,jj} - p^{-1}, \ldots, y_{jp,pj}, y_{j1,2j}, y_{j1,3j}, \ldots, y_{jp-1,pj})'$$

and each $<\underline{y}_{j\ell}>$, $1 \le j < \ell \le m$, is a p^2-vector

$(y_{j1,1\ell}, \ldots y_{jp,p\ell}, y_{j1,2\ell}, y_{j2,1\ell}, \ldots, y_{jp-1,p\ell}, y_{jp,p-1\ell})'$. Using Anderson and Stephens (1972, p.616) it follows that $n^{\frac{1}{2}}\underline{y}$ has covariance matrix $\Sigma = [(p-1)(p+2)]^{-1}$ blockdiag$((2C_{mp} \otimes \text{blockdiag}(C_{pp}, 2I_{\binom{p}{2}})),$

$$I_{\binom{m}{2}} \otimes \text{blockdiag}(C_{pp}, p^{-1} I_{\binom{p}{2}} \otimes D))$$

where $C_{mp} = I_m - p^{-1} 1_{mm}$, 1_{mm} is an $m \times m$ matrix of ones, and $D = \begin{bmatrix} p+1, & -1 \\ -1, & p+1 \end{bmatrix}$.

Since $C_{mp}^{-1} = C_{m,m-p}$ and $C_{pp}^{-} = I_p$ is a generalised inverse of C_{pp}, it follows that

$$\Sigma^{-} = \text{blockdiag}((\tfrac{1}{4}(p-1)(p+2)C_{m,m-p} \otimes \text{blockdiag}(I_p, 2I_{\binom{p}{2}})),$$

$$I_{\binom{m}{2}} \otimes \text{blockdiag}((p-1)(p+2)\, I_p, I_{\binom{p}{2}} \otimes D_1))$$

is a generalised inverse of Σ, where $D_1 = \begin{bmatrix} (p^2-1), & (p-1) \\ (p-1), & (p^2-1) \end{bmatrix}$. As Σ has rank $\nu(m,p)$ and the quadratic form $n\, \underline{y}'\Sigma^{-} \underline{y}$ reduces to B_{mp} as above, it follows that B_{mp} is asymptotically distributed as χ^2 on $\nu(m,p)$ degrees of freedom.

The only change necessary when $m = p$ is that $C_{m,m-p}$ should be replaced by I_p $(= C_{pp}^{-})$. We obtain B_{pp} as above, where now Σ has rank $\nu(p,p)$.

Remarks 1. When $p \geq 4$, B_{pp} is precisely the 'Bingham' statistic obtained by Prentice (1981, (3.4)) from consideration of the $(\underset{\sim}{0},2)$-th and $(\underset{\sim}{0},1,1)$-th characters of the irreducible continuous representations of the rotation group $O^{+}(p)$.

2. When $m = 1$, B_{1p} is the conventional Bingham statistic for directional data, since then $Z_2 = \underset{\sim}{0}_{pp}$ and Z_1 is symmetric.

4. Axial orientation statistics

The distribution (3.2) (or (2.2)) may be specialised to give a parametric model suitable for X-shapes. We require a probability density on $O(m,p)$ invariant under sign changes of any row of the random variable X. This is achieved if A satisfies the conditions $a_{jk\ell q} = 0$ if $j \neq q$. We write $B^{(ax)}(X,m,p,E)$ for the density

$$\mathrm{etr}(-\gamma P^{-1} I_p - \tfrac{1}{2} \sum_{j=1}^{m} E_j \underset{\sim}{x}_j \underset{\sim}{x}_j') \tag{4.1}$$

where $\gamma = \gamma(E)$ is a normalising constant, E is an $m \times p \times p$ array with jth layer E_j, $1 \leq j \leq m$, $e_{jk\ell} = a_{jk,\ell j}$, $1 \leq k,\ell \leq p$, and $\underset{\sim}{x}_j$ is the j-th row of the random variable X. Since $B^{(ax)}(X,m,p,E) = B^{(ax)}(X,m,p,E+E^{*})$ for all real $m \times p \times p$ arrays with elements $e^{*}_{jk\ell} = e_j$, dependent on layer only, it follows that the distribution (4.1) has parameter space of dimension $\eta(m,p) = m\binom{p+1}{2} - m = \tfrac{1}{2} m(p-1)(p+2)$, when $m < p$. If $m = p$, $e^{*}_{jk\ell}$ may be of the more general form $e_j + f_{k\ell}$ where $F = (f_{k\ell})$ is $p \times p$ symmetric, and without loss of generality, $f_{pp} = 0$. Hence for square axial orientations the parameter space of the distribution (4.1) has dimension $\eta(p,p) = p\binom{p+1}{2} - \binom{p+1}{2} - p + 1 = \tfrac{1}{2}(p-1)^2(p+2) = \eta(p-1,p)$. Where convenient to identify parameters uniquely, we shall assume that

$$e_{jpp} = 0 \quad \text{for all } j, \ 1 \le j \le m, \text{ and if } m = p, \ E_p = 0 \text{ also}. \tag{4.2}$$

A random sample from the distribution (4.1) has log likelihood

$$-n\gamma(E) - \tfrac{1}{2}n \ \text{trace} \sum_{j=1}^{m} E_j Y_j \tag{4.3}$$

where $Y_j = n^{-1} \sum_{i=1}^{n} \underline{x}_j^{(i)} \underline{x}_j^{(i)'}$, $\underline{x}_j^{(i)}$ representing the jth row of X_i. The distinct elements of $Y_1, \ldots, Y_m$, excluding those corresponding to (4.2), are minimally sufficient for E. As in Section 3 we obtain

Theorem 4.1

For sufficiently large n, with probability 1 there exists a unique MLE $\hat{E}$ of E, subject to (4.2), which is the solution of

$$\left(\frac{\partial\gamma}{\partial E}\right)_{(E=\hat{E})} = -\tfrac{1}{2}(Y_1, \ldots, Y_m,) .$$

An axial Bingham statistic, suitable for testing uniformity against alternatives (4.1) may be obtained from a simplified version of Theorem 3.2. From consideration of the asymptotic null distribution of the first $m\binom{p+1}{2}$ elements of $n^{\frac{1}{2}}\underline{Y}$ we obtain

Theorem 4.2

On the null hypothesis of uniformity the statistic $B_{mp}^{(ax)} = \tfrac{1}{2}n(p-1)(p+2)\ (\text{trace}(\sum_{j=1}^{m} Y_j^2) - m/p)$ is asymptotically distributed as χ^2 on $\eta(m,p)$ degrees of freedom.

Remark S_u (Theorem 2.2) $= \tfrac{1}{2}np(p-1)(p+2)\ (\text{trace}((\sum_{j=1}^{m} Y_j)^2) - m^2/p)/m(p-m)$, provided $m < p$.

5. Hybrids

A parametric model suitable for T-shapes may be obtained from the generalised Khatri-Mardia distribution

$$\operatorname{etr}(-\gamma(mp)^{-1} I_{mp} - \tfrac{1}{2}AX \otimes X' + A\mu \otimes X') \,, \qquad (5.1)$$

from (3,1), conditional on $XX' = I_m$, when $\mu \neq 0_{mp}$. We require that the pdf should be invariant under sign changes of any of the first m_1, rows (the axes) of X. This can be achieved by requiring that $\mu = \begin{pmatrix} 0_{m,p} \\ \mu_2 \end{pmatrix}$, where μ_2 is an $m_2 \times p$ matrix of means, $0 < m_2 = m - m_1 < m$, and that A satisfies the condition $a_{jk,\ell q} = 0$ if $j \neq q$ and $\min(j,q) \leq m_1$. We write $B^{(hy)}(X, m_1, m_2, p, \mu_2, E, A_2)$ for the density

$$\exp(-\gamma - \tfrac{1}{2}\operatorname{trace}(\sum_{j=1}^{m_1} E_j x_j x_j') + \operatorname{trace}(A_2\mu_2 \otimes X'_{(2)} - \tfrac{1}{2} A_2 X_{(2)} \otimes X'_{(2)})) \qquad (5.2)$$

where $\gamma = \gamma(E, A_2)$ is a normalising constant, E is as in Section 4, but with only m_1 layers, A_2 is the $m_2p \times m_2p$ submatrix of A corresponding to $X_{(2)}$, the last m_2 rows of the random variable X. Since

$$B^{(hy)}(X, m_1, m_2, p, \mu_2, E, A_2) = B^{(hy)}(X, m_1, m_2, p, \mu_2, E + E^*, A_2 + A_2^* \otimes I_p)$$

for all real E^* as in Section 4 (but with only m_1 layers), and all real symmetric $m_2 \times m_2$ matrices A_2^*, it follows that when $m = m_1 + m_2 < p$, the distribution (5.2) has parameter space of dimension

$$\phi(m_1, m_2, p) = m_1\binom{p+1}{2} + \binom{m_2p+1}{2} - m_1 - \binom{m_2+1}{2} + m_2p = \eta(m_1,p) + \nu(m_2,p) + m_2p \,.$$

When $m = p$, the more general result

$$B^{(hy)}(X, m_1, m_2, p, \mu_2, E, A_2) = B^{(hy)}(X, m_1, m_2, p, \mu_2, E + E^*, A_2 + (A_2^* \otimes I_p) + (I_{m_2} \otimes F))$$

obtains, where F and E^* are of the more general form in Section 4, and so, the distribution (5.2) has parameter space of dimension

$$\phi(m_1', m_2, p) = \eta(m_1', p) + \nu(m_2, p) + m_2p \,, \text{ where } m_1' = \min(m_1, p - m_2 - 1).$$

Where convenient to identify parameters, we shall assume that A_2 satisfies (3.3), and that

$$e_{jpp} = 0 \text{ for all } j,\ 1 \leq j \leq m_1, \text{ and if } m_1 + m_2 = p,\ E_{m_1} = 0 \qquad (5.3)$$

A random sample from the distribution (5.2) has log likelihood

$$-n\gamma(E, A_2) - \tfrac{1}{2}n \operatorname{trace}\left(\sum_{j=1}^{m_1} E_j Y_j\right) + n \operatorname{trace}\left(A_2 \mu_2 \otimes \bar{X}'_{(2)} - \tfrac{1}{2} A_2 Y_{(2)}\right) \tag{5.4}$$

where $\bar{X}_{(2)}$ is the $m_2 \times p$ matrix of means of the last m_2 rows of the data matrices $X_1, \ldots, X_n$, and $Y_{(2)}$ is the $m_2 p \times m_2 p$ submatrix of Y corresponding to the last m_2 rows. The distinct elements of $\bar{X}_{(2)}, Y_1, \ldots Y_{m_1}, Y_{(2)}$, excluding those corresponding to (3.3) and (5.3), are minimally sufficient for μ_2, E and A_2.

<u>Theorem 5.1</u>

For sufficiently large n, with probability 1 there exists a MLE $\hat{\theta} = (\hat{\mu}_2, \hat{A}_2, \hat{E})$ of $\theta = (\mu_2, A_2, E)$, subject to (3.3) and (5.3), which is the solution of

$$\left(\frac{\partial\gamma}{\partial\mu_{2,j\ell}}\right)_{(\theta=\hat{\theta})} = \sum_{kq} a_{jk,\ell q}\, \bar{x}_{2,qk}$$

$$\left(\frac{\partial\gamma}{\partial E}\right)_{(\theta=\hat{\theta})} = -\tfrac{1}{2}\,(Y_1, \ldots, Y_{m_1'})$$

$$\left(\frac{\partial\gamma}{\partial A_2}\right)_{(\theta=\hat{\theta})} = \mu_2 \otimes \bar{X}'_{(2)} - \tfrac{1}{2}\, Y_{(2)} \quad .$$

A large sample test of uniformity of T-shapes against alternatives (5.2) may be obtained by slight modification of Theorems (3.2) and (4.2). We obtain a hybrid Rayleigh-Bingham statistic:-

<u>Theorem 5.2</u>

On the null hypothesis of uniformity, the statistic $B^{(hy)}_{m_1 m_2 p} = B^{(ax)}_{m_1 p} + B_{m_2 p} + pn \operatorname{trace} \bar{X}_{(2)}\bar{X}'_{(2)}$, with $B^{(ax)}_{m_1 p}$ as in Theorem 4.2, and $B_{m_2 p}$ as in Theorem 3.2, is asymptotically distributed as χ^2 on $\phi(m_1', m_2, p)$ degrees of freedom.

Proof

$\bar{X}_2$ has expectation 0_{m_2p} and is asymptotically uncorrelated with $Y_1, \ldots, Y_{m_1}$ and $Y_{(2)}$. Using the method of Theorem 3.2 we obtain a quadratic form in $m_2p + \binom{m_2p+1}{2} + m_1\binom{p+1}{2}$ variables with covariance matrix of rank $\phi(m_1', m_2, p)$, which reduces to $B^{(hy)}_{m_1m_2p}$ as stated.

6. Exponential models and regression estimators

It is instructive to consider these matrix Bingham and Khatri-Mardia distributions within the context of the rotationally invariant exponential family of distributions for orientation statistics, obtained by generalising Beran's (1979) exponential family for directional data. Provided good multivariate density estimates are available, the obvious analogues of Beran's regression estimators ((1.10) ibid.) and goodness of fit tests (Section 5, ibid.) should be considerably more convenient computationally than exact MLEs and likelihood ratio tests.

Consider first the case $m = p$, and the general exponential model $\exp(h(X) - \gamma(h))$, $h \in \oplus_g M_g$, where M_g is associated with the g-th character of the irreducible continuous representations of $0^+(p)$, as in Prentice (1981). The von Mises-Fisher matrix distribution spans $M_{\underset{\sim}{0},1}$, of dimension p^2, and has basis $\beta_1 = \{x_{ij}, 1 \leq i,j \leq p\}$. The generalised Bingham matrix distribution (3.2) spans $M_2 = M_{\underset{\sim}{0},2} \oplus M_{\underset{\sim}{0},1,1}$ of dimension $\nu(p,p)$ when $p \geq 4$, and spans $M_2 \oplus M_1$ when $p = 3$. A basis is provided by

$\beta_2 = \{x_{ij}x_{k\ell}, 1 \leq i \leq k \leq p, 1 \leq j \leq \ell \leq p$, excluding cases corresponding to (3.3)$\}$.

When $m < p$, corresponding results are obtained by excluding the last $(p-m)$ rows of X. We obtain M_2 of dimension $\nu(m,p)$ with basis

$$\{x_{ij}x_{k\ell}, 1 \leq i \leq k \leq m, 1 \leq j \leq \ell \leq p, \text{ excluding cases (3.3)}\}.$$

For X-shapes, and the distribution (4.2) we proceed similarly. We obtain $\{x_{ij}x_{ik},\ 1 \le i \le m',\ 1 \le j \le k \le p$, excluding cases corresponding to (4.2)$\}$, of dimension $\eta(m',p)$. For hybrids, and the distribution (5.2), a basis is $\{x_{ij},\ m_1 < i \le m,\ 1 \le j \le p\} \oplus \{x_{ij}x_{ik},\ 1 \le i \le m_1',\ 1 \le j \le k \le p$ excluding cases corresponding to (5.3)$\} \oplus \{x_{ij}x_{k\ell},\ m_1 < i \le k \le m,\ 1 \le j \le \ell \le p$, excluding cases corresponding to (3.3)$\}$, of dimension $\phi(m_1',m_2,p)$.

Since Beran's results (1979) apply to the canonical exponential family of any compact space, his formulae (1.10) for estimators, (5.5) for approximate tests, and their extensions to interval estimation, may be used in large samples on orientation statistics of all types, provided only that suitable robust multivariate density estimates are available.

References

ANDERSON, T.W. and STEPHENS, M.A. (1972). Tests for randomness of directions against equatorial and bimodal alternatives. Biometrika 59, 613-621.

BERAN, R. (1979). Exponential models for directional data. Ann. Statist. 6, 1162-1178.

BERK, R.H. (1972). Consistency and asymptotic normality of MLEs for exponential models. Ann. Math. Statist. 43, 193-204.

BINGHAM, C. (1974). An antipodally symmetric distribution on the sphere. Ann. Statist. 2, 1201-1225.

BINGHAM, C. (1976). Expansions related to a hypergeometric function arising in an antipodally symmetric distribution on the sphere. Directional Data Project, Research Report 11, Dept. Stats. Leeds University.

DOWNS, T.D. (1972). Orientation statistics. Biometrika 59, 665-676.

JUPP, P.E. and MARDIA, K.V. (1979). Maximum likelihood estimators for the matrix von Mises-Fisher and Bingham distributions. Ann. Statist. 7, 599-606.

KHATRI, C.G. and MARDIA, K.V. (1975). The von Mises-Fisher matrix distribution. Directional Data Project, Research Report 1, Dept. Stats. Leeds University.

KHATRI, C.G. and MARDIA, K.V. (1977). The von Mises-Fisher matrix distribution in orientation statistics. J. Roy. Statist. Soc. Ser. B, 39, 95-106.

MARDIA, K.V. and KHATRI, C.G. (1977). Uniform distribution on a Stiefel manifold. J. Mult. Analysis 7, 468-473.

PRENTICE, M.J. (1981). Testing for uniformity of orientation statistics. Submitted to J. Mult. Analysis.

SRIVASTAVA, M.S. and CARTER, E.M. (1980). Asymptotic expansions for hypergeometric functions. In Multivariate Analysis V (P.R. Krishnaiah, ed.) North Holland Publishing Company.

THEOBALD, C.M. (1975). An inequality with application to multivariate analysis. Biometrika 62, 461-466.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER T.R. #192, Series 2	2. GOVT ACCESSION NO. AD-A097 460	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ANTIPODALLY SYMMETRIC DISTRIBUTIONS FOR ORIENTATION STATISTICS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Michael J. Prentice		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0322
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Princeton University Princeton, N.J. 08540		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 436) Arlington, Virginia 22217		12. REPORT DATE
		13. NUMBER OF PAGES 15
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Orientation statistics, Bingham matrix distribution, antipodal symmetry, exponential family.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The conventional antipodally symmetric Bingham matrix distribution on the Stiefel manifold is generalised. Large sample maximum likelihood estimation and uniformity tests are discussed, and a parametric model for axial orientations (X-shapes) is suggested. A generalisation of the Khatri-Mardia matrix distribution is developed to provide a model suitable for hybrids (T-shapes). Beran's results on exponential models for directional data are extended to orientation statistics to provide regression estimators and goodness-of-fit tests as alternatives to maximum likelihood estimation and likelihood ratio tests.